EQUIVALENCY OF SENTENCE INTELLIGIBILITY LISTS FOR AUDIOLOGICAL ASSESSMENT OF NAVAL PERSONNEL AND NAVY VOICE COMMUNICATION SYSTEMS

by

Thomas G. Giolas
Joseph R. Duffy
Robert J. Duffy
Harry S. Cooker
and
Russell L. Sergeant

NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY
REPORT NUMBER 759

Bureau of Medicine and Surgery, Navy Department
Research Work Unit M4305.08-3003DAC9.04

Transmitted by:

J. Donald Harris

J. Donald Harris, Ph.D.
Head, Auditory Research Branch

Reviewed and Approved by:

Charles F. Gell

Charles F. Gell, M.D., D.Sc.(Med)
SCIENTIFIC DIRECTOR
NavSubMedRschLab

Approved and Released by:

R. L. Sphar

R. L. Sphar, CDR MC USN
OFFICER IN CHARGE
NavSubMedRschLab

SUMMARY PAGE

THE PROBLEM

To determine the equivalence of ten Central Institute of the Deaf (CID) Sentence Intelligibility Lists and relate that information to the equivalence of corresponding revisions by the Naval Submarine Medical Research Laboratory (R-CID) of the ten CID Lists.

FINDINGS

(1) Two five-list combinations of the CID Lists yielded equivalent scores.

(2) Seven R-CID Lists yielded equivalent scores.

(3) The mean scores for eight R-CID Lists were significantly lower than scores obtained with the corresponding CID Lists.

(4) Both the CID and the R-CID sets of Lists contain a sufficient number of equivalent lists to render them useful in Navy research and clinical settings.

(5) The R-CID lists have an advantage of more equivalent lists and greater sensitivity to frequency distortion.

APPLICATION

Sentence Intelligibility lists can be used to provide oto-audiologists additional evaluation of various hearing problems incurred by Navy personnel. Also, the lists can be used by communications engineers involved in determining figure-of-merit for Navy systems, such as in the case of message sending from hyperbaric or underwater environments which create distortions in speech.

ADMINISTRATIVE INFORMATION

This investigation was conducted as part of Bureau of Medicine and Surgery Work Unit M4305.08-3003DAC9, Development of Auditory Screening Standards for Submarine/Shipboard Personnel. The assistance of four of the five investigators was furnished under ONR Contract with the University of Connecticut (N00014-67-A-0197-0001). The present report is Number 4 on this work unit. It was submitted for review on 10 July 1973, approved for publication on 14 Sept. 1973 and designated as NavSubMedRschLab Report No. 759.

PUBLISHED BY THE NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY

ABSTRACT

The purpose of this study was to investigate the equivalency of several sentence intelligibility lists for use in clinical hearing tests of Navy personnel and research facilities requiring figure-of-merit indices of voice communication systems. Ten Central Institute for the Deaf (CID) Sentence Intelligibility Lists and revisions of those ten lists (R-CID) were presented to two groups of 30 listeners each. Scores were the number of correct key words identified for each sentence list. The following results were obtained: (1) Two five-list combinations of the CID Sentence Lists (A, D, E, G, J and B, D, E, G, J) yielded equivalent scores. (2) Seven R-CID Sentence Lists (A, C, E, F, G, I, J) yielded equivalent scores. (3) The mean scores for eight R-CID Sentence Lists (A, B, C, D, E, F, G, J) were significantly lower than scores obtained with the corresponding CID Lists. (4) The results suggest that both sets of sentence lists contain a sufficient number of equivalent lists to render them useful in most Navy research and clinical settings. The R-CID Lists have a slight advantage in that they contain more equivalent lists and appear to be more sensitive to frequency distortion.

EQUIVALENCY OF SENTENCE INTELLIGIBILITY LISTS FOR AUDIOLOGICAL ASSESSMENT OF NAVAL PERSONNEL AND NAVY VOICE COMMUNICATION SYSTEMS

INTRODUCTION

Sentences have re-emerged as a desirable message for the measurement of speech intelligibility (Berger,[1] Giolas,[4] Harris, et al.,[7] and Jerger, et al.[10]). It is argued that they present a more natural listening task than do the widely used monosyllabic word lists (Hirsh, et al.[8]) and take advantage of crucial parameters utilized in understanding connected speech. Lists of Central Institute for the Deaf (CID) Sentences (Silverman and Hirsh[11]) have been suggested as one possible message for clinical and research use. This series of sentence lists was developed to be representative of colloquial speech, and a close relationship was found between scores obtained using selected lists of the series and a sample of continuous discourse (Giolas[4]). The sentence lists in this selected set are available in written form (Davis and Silverman[2]) and are easy to administer and score. In addition, lists in that set were revised by Harris, et al.[7] to provide a greater homogeneity of sentence length while attempting to maintain the colloquial speech criterion. This revised set of lists is known as the Revised CID Sentence Lists (R-CID).

Little information is available on the equivalency of speech intelligibility scores obtain with either the CID or R-CID Sentence Lists. Such list equivalency information is vital if these lists are to be useful in the research and clinical setting. Therefore, the purpose of this study was to investigate the equivalency of speech intelligibility scores obtained with the two sets of sentence intelligibility lists, CID and R-CID. In addition, the effect of the sentence modification procedure used to form the set of R-CID Lists was analyzed.

METHOD

Filtering and Recording

Each set (CID and R-CID) of sentence intelligibility lists consists of ten lists, each list having 50 key words embedded in ten sentences. All lists were read by an adult male* and recorded using a Shure 546 Microphone and an Ampex 300-2 Tape Recorder. Voice intensity was monitored with a VU-meter during the recording sessions.

To produce the error responses necessary for comparisons among the test lists, both sets of sentence intelligibility lists were distorted by 420 Hz low-pass filtering using an Allison 2B filter between two Ampex 300-2 tape recorders. Previous research indicates that this distortion level facilitates comparisons (Duffy and Giolas[3]). The VU meters of the recorders were matched and use of a calibration tone ensured appropriate playback levels. Five practice sentences preceded each experimental tape to familiarize the subjects with the listening task.

**The voice on all these recordings was that of the second-named author, J. R. Duffy.*

Sentence Presentation

Sixty normal-hearing adult males were divided into two groups of thirty each. One group heard all ten CID Lists, the other all ten R-CID Lists. Each group was further divided into three sub-groups of ten listeners who heard a different randomization of their ten lists. Test tapes were played on an Ampex PR-10 recorder, via an Altec 1569A amplifier to 49 TDH-39 earphones mounted in Otocups. The test room was specifically designed for listening experiments. Each sub-group responded to a monaural presentation of 10 sentence intelligibility lists presented at a comfortable loudness level and administered over two sessions. Subjects were asked to write down, word for word, what they heard after each sentence was presented. They were encouraged to respond with complete sentences even if this required guessing at words or phrases which they did not understand. Scores were based on the number of 50 key words in each list correctly identified.

RESULTS AND ANALYSIS

Data were subjected to an analysis of variance for a two factor experiment (A, CID/R-CID; B, Lists) with repeated measures on one factor (Winer[12]). The results of this initial analysis are summarized in Table 1. The significant "F" (.01) obtained for the A effect suggests overall differences between the two sets of sentence lists for the distortion of 420 Hz low-pass filtering. The

Table 1. Summary of a Two Factor With Repeated Measures for One Factor Analysis of Variance (Winer, 1962, p. 306)

Source of Variation	d.f.	SS	MS	F	Required F (.01)
Between S's.	59	22919.4183			
A (CID & R-CID Lists)	1	12303.4817	12303.4817	67.2199*	7.10
Subject within groups	58	10615.9366	183.0334		
Within S's.	540	24584.3000			
B (Lists A-J)	9	2895.7017	321.7447	8.8384*	2.41
AB	9	2686.0349	298.4484	8.1984*	2.41
Bx S's within groups	522	19002.5634	36.4034		

significant "F" (.01) obtained for the B effect (Lists A-J) further suggests differences between some lists within each of the two sets. The significant (.01) AB interaction between the two sets of sentence lists and individual lists indicates additional analysis was appropriate to evaluate separately lists within each set.

Equivalency of CID Sentence Lists

Means and standard deviations of scores obtained for the ten sentence lists comprising the set of CID Sentence Lists are summarized in Table 2. The significant "F" (.01) computed separately for these sentence lists indicate differences between mean scores for lists comprising the CID set. The differences between the means for all combinations of the ten sentence lists are tabulated in Table 3, along with an indication of which differences were found to be statistically significant employing a "critical difference" procedure outlined by Lindquist[9]. Mean scores obtained were not significantly (.01) different for lists A, D, E, G, J or for

Table 2. Summary of Mean Scores and Standard Deviations Obtained with the CID Sentence Lists (Scores based on number of 50 key-words correctly identified)

Sentence List	Mean Score	S.d.
A	36.93	7.05
B	40.43	7.87
C	41.83	5.21
D	37.10	6.76
E	39.73	6.54
F	32.93	5.63
G	40.30	6.12
H	33.67	6.25
I	30.93	8.96
J	38.30	6.76

Table 3. Differences Between Mean Scores for CID Sentence Lists

Lists	A	B	C	D	E	F	G	H	I	J
A		3.50*	4.90*	.17	2.80	4.00*	3.37	3.27	6.00*	1.37
B			1.40	3.33	.70	7.50*	.13	6.77*	9.50*	2.13
C				4.73*	2.10	8.90*	1.53	8.17*	10.90*	3.53*
D					2.63	4.17*	3.20	3.44	6.17*	1.20
E						6.80*	.57	6.07*	8.80*	1.43
F							7.37*	.73	2.00	5.37*
G								6.64*	9.37*	2.00
H									2.73	4.64*
I										7.37*
J										

*Significant at .01 level

lists B, D, E, G, and J. The mean intelligibility scores for the lists comprising either of these list combinations did not differ by more than 7%. It was concluded that either five-list combination yields equivalent speech intelligibility scores when presented under the level of distortion employed in this study. A further analysis of Table 3 yielded the more detailed list equivalency breakdown tabulated in Table 4. In situations where equivalent lists are necessary, care should be taken to employ only these lists which yielded statistically similar scores.

Equivalency of Revised CID Sentence Lists

Means and standard deviations of scores obtained with the R-CID set of ten lists are summarized in Table 5. The significant "F" (.01) computed separately for these sentence lists indicated differences among mean scores for the lists within this set. The differences between the means for all combinations of the ten sentence lists are tabulated in Table 6, along with an indication of which differences were statistically significant (.01). Mean scores

Table 4. Combination of CID and Revised CID Lists Considered to Yield Equivalent Scores

C.I.D. Lists	Rev. C.I.D. Lists
B=D=E=G=J	A=C=E=F=G=I=J
A=D=E=G=J	B=F=G=I=J
B=C=E=G	
F=H=I	
H=I	H=J

Table 5. Summary of Mean Scores and Standard Deviations Obtained with the Revised CID Sentence Lists

Sentence List	Mean Score	S.d.
A	30.33	8.02
B	26.63	8.62
C	29.87	5.65
D	21.73	8.00
E	30.00	5.80
F	27.70	7.32
G	27.97	6.89
H	31.47	5.91
I	27.27	7.46
J	28.63	8.31

Table 6. Differences Between All Combinations of Means Obtained for Lists Comprising the Revised CID Sentence Lists

Lists	A	B	C	D	E	F	G	H	I	J
A		3.70*	.4667	8.6000*	.3333	2.6333	2.3667	1.1333	3.0667	1.7000
B			3.2333*	4.9000*	3.3667*	1.0667	1.3333	4.8333*	.9333	2.0
C				8.1333*	.1334	2.1666	1.9000	1.6000	2.6000	1.2333
D					8.2667*	5.9667*	6.2333*	9.7333*	5.9327*	6.9000*
E						2.3000	2.0334	1.4666	2.7334	1.3667
F							.2666	3.7666*	.4333	.9333
G								3.5000*	.7000	.6667
H									4.20*	2.8333
I										1.3667
J										

*Significant at the .01 level

obtained for lists A, C, E, F, G, I, and J were not found to differ significantly. The mean intelligibility scores for these lists did not vary by more than 8.40%. It was concluded that these seven lists yield equivalent speech intelligibility scores when presented with the filtering distortion employed in this study. Further analysis of Table 6 yields the more detailed list equivalency breakdown tabulated in the right column of Table 4.

It is interesting to note that while seven of the ten lists comprising the R-CID Lists were found to produce equivalent speech intelligibility scores, only five of ten original CID Lists were statistically equivalent.

Effects of Sentence Modification

The final analysis consisted of investigating the effect of sentence modification. The modification procedure involved both the addition and elimination of words from individual sentences comprising the original lists, in order

to more closely equate them with regard to sentence length (Harris, et al.[7]). In many cases, the new sentence list had an altered meaning from its original counterpart. However, attempts were made to retain the colloquial nature of the original lists.

An "F" was computed between each original list and its corresponding revised list (Winer[12]). As can be seen in Table 7, eight of the original ten sentence lists were significantly altered by the sentence modification procedure. The mean scores were significantly lower for all R-CID Lists, except Lists H and I. However, the mean scores of even these two revised lists were lower than were the original two counterpart lists. The relationship between the CID and R-CID Lists is graphically illustrated in Figure 1. It is apparent that the revised lists (solid bars) are more sensitive to frequency distortion (420 Hz low pass) than are the original CID lists.

Table 7. Summary of Means and F Tests Computed For Each Original CID Sentence List and Its Corresponding Revised List

List	C.I.D.	Rev. C.I.D.	F
A	36.93	30.33	12.80*
B	40.43	26.63	55.94*
C	41.83	29.86	42.06*
D	37.10	21.73	69.36*
E	39.73	30.00	27.83*
F	32.93	27.70	8.04*
G	40.30	27.97	32.09*
H	33.66	31.47	1.42
I	30.93	27.27	3.95
J	38.30	28.63	27.45*

*Significant at the .01 level

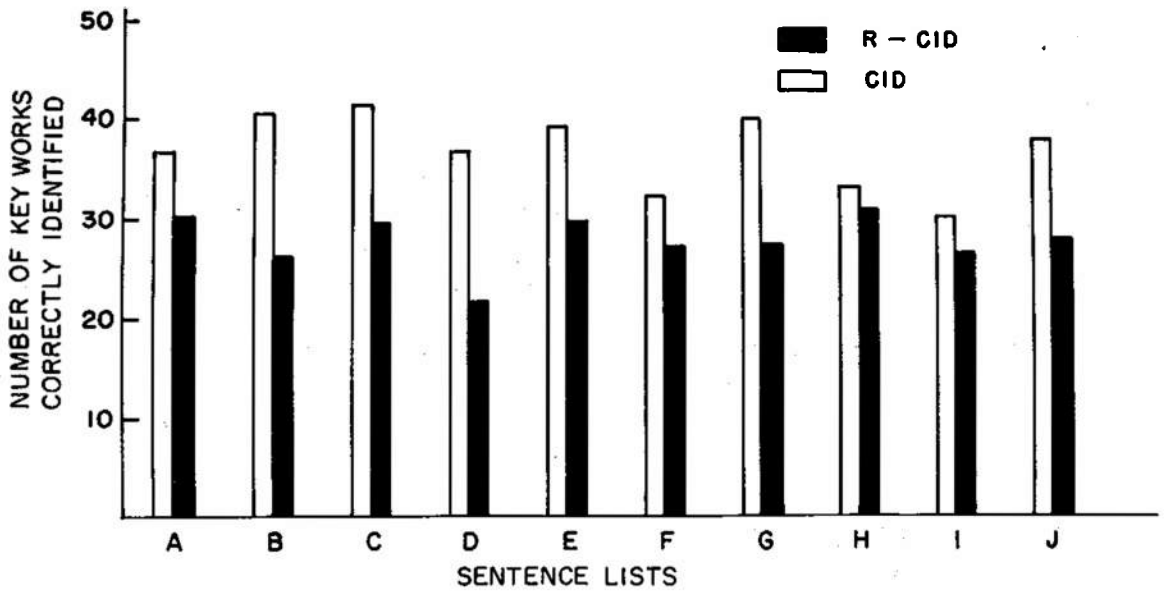

Fig. 1. Comparison of mean scores obtained for all CID Sentence Lists and their corresponding revised lists. Scores were based on number of key words correctly identified.

CONCLUSIONS

The Naval Submarine Medical Research Laboratory provides solutions for the Navy to problems in the area of hearing encountered by the Naval Submarine Force and by Navy personnel working within diving/swimming environments. To accomplish this mission, continued use is required of various indices of speech perception to evaluate the capability to hear speech by Navy personnel as well as to determine the figure-of-merit for Navy communication systems. Sentence intelligibility tests have certain advantages which render them preferable to other speech tests in certain clinical and research settings. The results of the present study indicate that both the CID and the R-CID sentence intelligibility tests contain enough equivalent lists to make them valuable testing tools. There appear to be advantages to using the R-CID Lists, since they include more equivalent lists and are more sensitive to frequency distortion. The R-CID Lists also seem less predictable than the CID Lists, suggesting that intelligibility scores obtained with the R-CID Lists would be less inflated by message prediction (Giolas, et al.[5]).

REFERENCES

1. Berger, K. A speech discrimination task using multiple-choice key words in sentences. J Aud Res, 9, 247-262, 1969.

2. Davis, M. and Silverman, S.R., (Ed.) Hearing and Deafness (Rev. ed.) New York: Holt, Rinehart and Winston, 1970.

3. Duffy, J.R. and Giolas, T.G. The effect of word predictability on sentence intelligibility. NSMRL Report No. 672, 1971.

4. Giolas, T.G. Comparative intelligibility scores of sentence lists and continuous discourse. J Aud Res 6, 31-38, 1966.

5. Giolas, T.G., Cooker, H.S. and Duffy, J.R. The predictability of words in sentences. J Aud Res VII, No. 1, 1972.

6. Harris, J.D., Haines, H.L., and Myers, C.K. The importance of hearing at 3 kc for understanding speeded speech. Laryngoscope 70, 131-146, 1960.

7. Harris, J.D., Haines, H.L., Kelsey, P.A. and Clack, T.D. The relation between speech intelligibility and electroacoustic characteristics of low fidelity circuitry. J Aud Res 1, 357-381, 1961.

8. Hirsh, I.J., Davis, H., Silverman, S.R., Eldert, E.G., and Benson, R.W. Development of materials for speech audiometry. J Speech Hear Disor 17, 321-337, 1952.

9. Lindquist, E.F. Design and Analysis of Experiments in Psychology and Education. Houghton Mifflin Company, Boston, Mass., 1953.

10. Jerger, J., Speaks, C. and Trammell, J.L. A new approach to speech audiometry. J Speech Hear Disor 33, 318-328, 1968.

11. Silverman, S.R. and Hirsh, I.J. Problems related to the use of speech in clinical audiometry. Ann Otol Rhinol and Laryngology 64, 1234-1244, 1955.

12. Winer, B.J. Statistical Principles in Experimental Design. McGraw-Hill Book Company, New York, N.Y., 1962.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
NAVAL SUBMARINE MEDICAL RESEARCH LABORATORY, Naval Submarine Medical Center

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE
Equivalency of Sentence Intelligibility Lists for Audiological Assessment of Naval Personnel and Navy Voice Communication Systems

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Interim report

5. AUTHOR(S) *(First name, middle initial, last name)*
Thomas G. GIOLAS, Joseph R. DUFFY, Robert J. DUFFY, Harry S. COOKER and Russell L. SERGEANT

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
14 September 1973	9	12

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. M4305.08-3003DAC9	NSMRL Report No. 759
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

10. DISTRIBUTION STATEMENT
Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Naval Submarine Medical Center Box 600 Naval Submarine Base Groton, Connecticut 06340

13. ABSTRACT

The purpose of this study was to investigate the equivalency of several sentence intelligibility lists for use in clinical hearing tests of Navy personnel and research facilities requiring figure-of-merit indices of voice communication systems. Ten Central Institute for the Deaf (CIF) Sentence Intelligibility Lists and revisions of those ten lists (R-CID) were presented to two groups of 30 listeners each. Scores were the number of correct key words identified for each sentence list. The following results were obtained: (1) Two five-list combinations of the CID Sentence Lists (A, D, E, G, J and B, D, E, G, J) yielded equivalent scores. (2) Seven R-CID Sentence Lists (A, C, E, F, G, I, J) yielded equivalent scores. (3) The mean scores for eight R-CID Sentence Lists (A, B, C, D, E, F, G, J) were significantly lower than scores obtained with the corresponding CID lists. (4) The results suggest that both sets of sentence lists contain a sufficient number of equivalent lists to render them useful in most Navy research and clinical settings. The R-CID lists have a slight advantage in that they contain more equivalent lists and appear to be more sensitive to frequency distortion.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0102-014-6600

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Speech Discrimination Testing						
Speech Reception						
Sentence Intelligibility						
Communications System Testing						
Audiological Speech Test						